Structure, Vol. 12, 1655–1667, September, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.str.2004.06.022

Novel Binding Site Identified in a Hybrid between Cholera Toxin and Heat-Labile Enterotoxin: 1.9 Å Crystal Structure Reveals the Details

Åsa Holmner,[1,*] Michael Lebens,[2] Susann Teneberg,[3] Jonas Ångström,[3] Mats Ökvist,[4] and Ute Krengel[1,5,*]
[1]Department of Chemistry and Bioscience
Chalmers University of Technology
P.O. Box 462
SE-40530 Göteborg
Sweden
[2]Institute of Medical Microbiology
Göteborg University
P.O. Box 435
SE-40530 Göteborg
Sweden
[3]Institute of Medical Biochemistry
Göteborg University
P.O. Box 440
SE-40530 Göteborg
Sweden
[4]Department of Biochemistry and Biophysics
Göteborg University
P.O. Box 462
SE-40530 Göteborg
Sweden
[5]Department of Chemistry
University of Oslo
P.O. Box 1033
Blindern, N-0315 Oslo
Norway

Summary

A hybrid between the B subunits of cholera toxin and *Escherichia coli* heat-labile enterotoxin has been described, which exhibits a novel binding specificity to blood group A and B type 2 determinants. In the present investigation, we have determined the crystal structure of this protein hybrid, termed LCTB*K*, in complex with the blood group A pentasaccharide GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ, confirming not only the novel binding specificity but also a distinct new oligosaccharide binding site. Binding studies revealed that the new specificity can be ascribed to a single mutation (S4N) introduced into the sequence of *Escherichia coli* heat-labile enterotoxin. At a resolution of 1.9 Å, the new binding site is resolved in excellent detail. Main features include a complex network of water molecules, which is well preserved by the parent toxins, and an unexpectedly modest contribution to binding by the critical residue Asn4, which interacts with the ligand only via a single water molecule.

Introduction

Carbohydrates are complex molecules with high potential for the encoding of biological information. Accordingly, their prominent role in disease mechanisms and recognition processes has vastly increased the interest in carbohydrates and carbohydrate binding proteins in recent years (Science, 2001; Dove, 2001). The diversity in the way that monosaccharide moieties can form complex structures far exceeds that of nucleotides and amino acids, which are limited to forming linear polymers. Consequently, a large variety of carbohydrate structures are found within the human body, which when combined with proteins and lipids on cell surfaces can be seen to clearly influence processes such as cell-to-cell communication, cellular differentiation, immune responses, and the progression of cancer (Sharon and Lis, 1993; Science, 2001). Furthermore, many microbes and specific toxins make use of surface-exposed carbohydrates to facilitate their attachment and subsequent entry into host cells (Beachey, 1981; Mirelman and Ofek, 1986; Karlsson, 1989). Proteins that recognize and bind carbohydrate structures are highly specialized and selective. Therefore, an understanding of carbohydrates and carbohydrate binding proteins is potentially important for the development of therapies against many severe human diseases (Maeder, 2002; Alper, 2001; Dove, 2001; Williams and Davies, 2001; Dennis, 2003).

The cholera toxin (CT) from *Vibrio cholerae* (Lai, 1980) and the heat-labile enterotoxin (LT) from enterotoxigenic *Escherichia coli* (Gyles, 1992) are structurally and functionally related proteins (Merritt and Hol, 1995). They consist of one enzymatically active A subunit (with ADP-ribosyl transferase activity) and five receptor binding B subunits. In the holotoxin, the B subunits are assembled into a doughnut-shaped pentamer with the A subunit noncovalently anchored in the center. Cholera toxin and *Escherichia coli* heat-labile enterotoxin primarily act as mediators of diarrheal diseases (i.e., cholera and the slightly milder form called "tourist or travelers diarrhea") (Holmgren, 1981). Upon infection, the organism secretes the toxin and the B-pentamer binds to receptors present on the target cells. The entire toxin is then internalized by a process that is not yet fully understood. In the cell, a cascade of reactions, triggered by the initial ADP-ribosylation of the ATPase Gsα, finally leads to elevated levels of cAMP, giving rise to a subsequent secretion of fluid and electrolytes from the cells.

An intriguing property of the toxins is their potent mucosal and systemic immunogenicity. The holotoxins are not only strong immunogens but also powerful mucosal adjuvants that stimulate immune responses to co-administered antigens (Simmons et al., 2001). The isolated B-pentamers are potent immunogens in their own right and, in fact, the anti-toxin antibody response upon exposure to the holotoxin is primarily directed toward the B subunits (Belisle et al., 1984; Holmes and Twiddy, 1983; Nashar et al., 1996). Furthermore, the B subunits have immune-modulating properties that are distinct from those of the holotoxins. Rather than acting as adjuvants, the B subunits have been shown to drive the induction of immunological tolerance to chemically and

*Correspondence: asa.holmner@chembio.chalmers.se (A.H.); ute.krengel@chembio.chalmers.se (U.K.)

genetically coupled antigens (Lebens et al., 2003; Sadeghi et al., 2002; Sun et al., 1994).

The CT and LT B-pentamers (CTB and LTB) share 83% sequence identity (corresponding to 17 substitutions among 103 amino acid residues). Both B-pentamers bind to the ganglioside GM1 (Galβ3GalNAcβ4(NeuAcα3)Galβ4GlcβCer) as their primary receptor (Holmgren, 1973; Holmgren et al., 1975). However, while CTB exclusively recognizes this one receptor, LTB has a broader specificity and has been shown to recognize glycoproteins and glycolipids with *N*-acetyllactosamine- (Galβ4-GlcNAc-) terminated carbohydrate chains, in addition to its primary receptor GM1 (Griffiths and Critchley, 1991; Holmgren et al., 1982, 1985; Karlsson et al., 1996; Orlandi et al., 1994; Teneberg et al., 1994).

In order to evaluate the influence of specific amino acid residues on carbohydrate binding specificity, a number of hybrids between recombinant CTB (rCTB) and LTB from human isolates (hLTB) were constructed (Bäckström et al., 1997; Ångström et al., 2000). These hybrids were generated by substituting amino acid residues of CTB for those at the corresponding positions in LTB. Purely by chance, a novel binding specificity (to type 2 blood group A and B antigens) was identified that, according to computer modeling studies, seemed unlikely to be linked to the GM1 binding site. The protein hybrid with the new binding specificity (termed LCTB*K*) exhibits a sequence midway between the parent toxins, differing from CTB in 8 residues and from hLTB in 9 residues, respectively. Furthermore, only one amino acid (the Asn residue at position 4) differs between this new protein hybrid and a hybrid with binding properties indistinguishable from hLTB, termed LCTB*H*, which exhibits a serine residue at this position. Molecular modeling lead to the proposal of a new binding site, distinct from the original GM1 binding site, in which the critical asparagine residue at position 4 plays a central role in anchoring the blood group A pentasaccharide to the protein (Ångström et al., 2000).

In the present study, we have solved the crystal structure of the protein hybrid LCTB*K* in complex with the blood-group A pentasaccharide GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ (Figure 1). Inspired by the knowledge that only one amino acid residue differs from a protein hybrid with LTB properties, we further extended the study by introducing the same point mutation in the recombinant hLTB molecule, thus exchanging serine to asparagine at position 4. This new hybrid, hLTB/S4N, was further investigated with respect to its binding properties by glycosphingolipid binding studies. Interestingly, the hLTB single mutant displays similar binding properties as the previously identified protein hybrid LCTB*K*, thus nailing down the new binding specificity to a single amino acid residue.

Results

The driving force for performing the present structural investigation was evidence for a novel binding site in the protein hybrid LCTB*K* based on glycosphingolipid binding studies and theoretical modeling (Ångström et al., 2000). The 1.9 Å crystal structure of the hybrid LCTB*K* in complex with the blood group A pentasaccharide GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ confirms the generation of a new binding specificity and the existence of a distinct new ligand binding site.

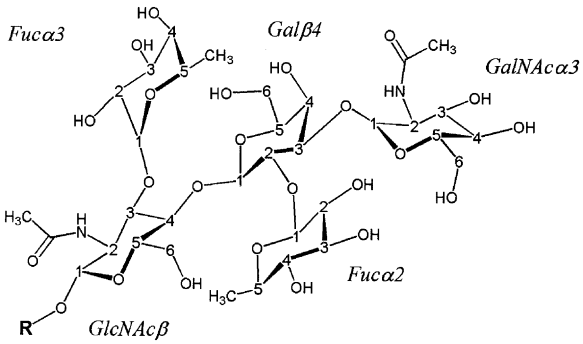

Figure 1. Schematic Drawing of the Blood-Group A Type 2 Pentasaccharide GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ

R stands for the core extension of the saccharide chain (R = 3Galβ4-GlcNAcβ3Galβ4Glcβ1Cer for A9-2 antigens). For comparison, in blood group B determinants, the GalNAc residue is replaced by Gal. In type 1 determinants, the β1-4 linkage between Gal and GlcNAc is replaced by a β1-3 linkage, respectively (and Fucα3 becomes β1-4 glycosidically linked to GlcNAc instead); thus, the terminal trisaccharide and the fucose unit are exchanged on GlcNAcβ (figure generated with ChemDraw Std.).

Final Crystallographic Model

The crystals contain two pentamers per asymmetric unit. Five pentasaccharides are bound to each pentamer. The three-dimensional fold of the parent proteins (Sixma et al., 1993; Merritt et al., 1994) is well preserved in the hybrid with each monomer containing a five-stranded antiparallel β sheet and two α helices that line the central pore of the toxin B-pentamer. The final crystallographic model (PDB entry 1TL0; R = 17.5%, R_{free} = 22.7%) contains 960 amino acid residues, 50 sugar residues, and 911 water molecules. Root-mean-square deviations (rmsd) from ideal geometry are 0.013 Å and 1.70° for bond lengths and angles, respectively (see Table 1). There is clear electron density for all sugar residues of the blood group A pentasaccharide (Figure 2A) as well as for the majority of the protein residues, except for a few solvent-exposed regions on the protein surface (the average real space correlation coefficient (r.s.c.c.) is 87%, based on a composite annealed OMIT map).

Novel Binding Site for Blood Group Antigens

Owing to differences in the local environment and the intrinsic flexibility of the pentasaccharide, the quality of the electron density differs slightly between the ten binding sites. Nevertheless, all five monosaccharide residues could be unmistakably modeled in all ten copies of the B subunit. The analysis of the carbohydrate binding site (as described below and detailed in Figure 2B and Table 2) is mainly based on the binding site at the interface of the D and H subunit of the B-pentamer, which is one of the highly solvent-exposed sites; however, the variation of ligand binding within the decamer is small (see Figure 3; rmsd of the C_α atoms of residues

Table 1. Data Collection and Refinement Statistics

X-Ray Source	Synchrotron	Rotating Anode
Space group	C2	C2
Unit cell parameters		
a (Å)	109.4	110.2
b (Å)	70.4	70.1
c (Å)	136.7	137.4
β (°)	91.9	92.9
Resolution (Å)	2.00	1.94
Completeness (%)	97.6 (96.8)	96.4 (93.0)
I/σ	18.1 (4.8)	19.6 (5.7)
R_{merge} (%)	4.9 (29.5)	5.6 (17.6)
Redundancy	4.3 (4.2)	5.4 (4.9)
Number of observed reflections	292,908 (38,364)	401,958 (30,795)
Number of unique reflections	68,619 (9,231)	74,812 (6,278)
Reflections in test set (%)	5	5
R_{value} (%)		17.5
R_{free} (%)		22.7
Rmsd bond lengths (Å)		0.013
Rmsd angles (°)		1.70
R.s.c.c.[a] (%)		
Protein		87
Carbohydrate		74
Water molecules		71
Average B factors (Å^2)		
Protein atoms		24.0
Carbohydrate atoms		33.7
Water molecules[b]		34.0
Ramachandran profile[c] (%)		
Most favorable		93.4
Additionally allowed		6.5
Generously allowed		0.1
Disallowed		0.0
Average solvent-accessible surface per binding site[d] (Å^2)		
Protein		343.7
Carbohydrate		457.6
PDB entry		1TL0

Values in parentheses correspond to the data in the highest resolution shell (2.1–2.0 Å and 2.0–1.9 Å for the synchrotron and rotating anode data sets, respectively).

[a] Calculated using the program CNS (Brünger et al., 1998) using a composite annealed OMIT map.

[b] The water molecules W1–W7 in the binding site have an average r.s.c.c. of 33.8%.

[c] According to the program PROCHECK (Laskowski et al., 1993).

[d] Calculated with the program AREAIMOL from the CCP4 program suite (CCP4, 1994). The calculation was performed on the protein complex including water molecules W1–W7.

involved in binding are 0.2 ± 0.1 Å, thus within the precision limits of the structure).

The blood group A binding site is positioned in a shallow cavity at the interface of two B subunits (see Figure 4), in a position roughly corresponding to the model proposed by Ångström et al. (Ångström et al., 2000). This new binding site is clearly distinct from the original GM1 binding site (Figure 4A). One of the two B subunits provides the majority of the amino acid contacts involved in direct carbohydrate protein interactions. The three terminal carbohydrate moieties GalNAcα3, Galβ4, and Fucα2 are rather firmly bound to the protein by direct hydrogen bonds to four protein residues, Gly45, Thr47, Asn94, and Gln3# from the adjacent B subunit (# indicating the adjacent B subunit). The Fucα2 residue is deeply buried in the otherwise shallow pocket and serves as an anchor for the entire oligosaccharide by binding to three of these four residues (Figures 2B and 4; Tables 1 and 2). In addition, its methyl group is involved in strong hydrophobic interactions with CE1 (or CE2) and CZ of Phe48. This residue alone contributes to 40% of the buried solvent-accessible surface area of the pentasaccharide (in average 164.6 of 457.6 Å^2; see Table 1).

Apart from the above-mentioned direct interactions, saccharide binding to LCTB*K* is mainly promoted indirectly, via water molecules. In fact, the majority of protein-ligand interactions in this system are of indirect nature. The water molecules cluster in two distinct networks (W4-7 and W1-3; see Figures 2B and 3). The former water cluster serves as an anchor for the reducing end of the pentasaccharide (GlcNAcβ), and for Fucα2, by mediating contacts to protein residues Gln16, Tyr18, Thr92, and Asn94. The latter water network links the terminal GalNAcα residue to Thr47, Gln3#, Asn4#, and Glu7#. It should be noted that Asn4, the residue distinguishing hybrid LCTB*K* from LCTB*H* (and instrumental for creating the novel binding specificity, as shown below), is part of the latter interaction network (for details, see Table 2). In contrast to our expectations, this residue exhibits no direct interactions with the saccharide ligand. Its only contact with the blood group A pentasaccharide is a single water-mediated hydrogen bond to GalNAcα3 3-OH.

The fifth monosaccharide unit of the blood group A pentasaccharide, Fucα3, is highly solvent exposed and does not form any hydrogen bonding interactions with the protein. Nevertheless, its electron density is clearly interpretable, due to extensive hydrophobic interactions of its methyl group with the peptide stretching from Gly45 CA/C to Ala46 CA/CB and due to additional internal stabilization within the pentasaccharide (for details, see Table 2 and Figure 2B).

Structural Comparison of LCTBK with Parent B-Pentamers

A comprehensive comparison with ten high-resolution CTB and LTB structures (60 B subunits in total) from the PDB (Berman et al., 2000) (see Experimental Procedures) indicates that the overall protein fold is very well preserved in LCTB*K* (rmsd for C_α atoms is 0.4 ± 0.1 Å), even in the novel binding site (see Figure 3). Largest differences are found for the loop participating in GM1 recognition (residues 51–59), which adopts an unusual conformation in one of the ten LCTB*K* subunits (subunit P). The local environment in this subunit, however, strongly suggests that this uncommon fold is induced by crystal contacts and hence can be considered an artifact introduced by crystallization, with no biological significance. The only relevant difference concerns a small part of the blood group A binding site, i.e., the tip of a loop, composed of residues 44–46, where the carbohydrate-protein interactions have induced a change in conformation. This is the region involved in strong hydrophobic interactions with the Fucα3 methyl group. Gly45, whose carbonyl group is engaged in two hydrogen bonds—with Galβ4 O4 and with GalNAcα2 N2—has moved by about 1.0 ± 0.5 Å to facilitate binding to

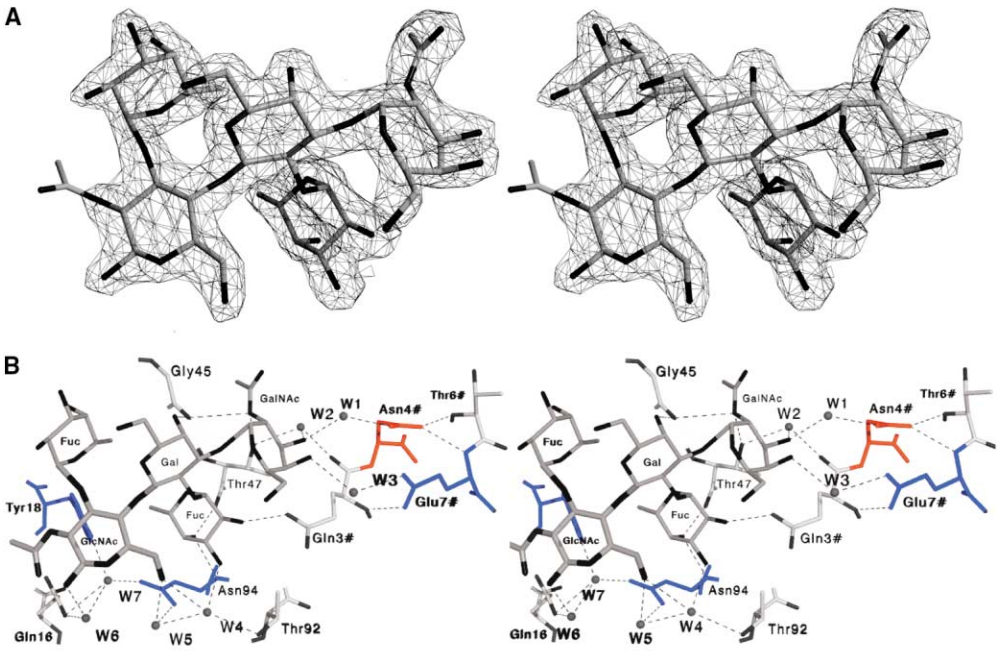

Figure 2. Stereo Pictures of the Pentasaccharide Ligand, Based on the Binding Site Involving Subunits D and H

(A) Exemplary ($2F_o - F_c$) electron density for the saccharide ligand, represented as chicken wire.

(B) Detailed stereo picture of the blood group A binding site. Note the modest role of the CTB-specific residue Asn4 (in red). This residue has been shown to be of critical importance for creating the novel binding specificity and is nevertheless bound to the pentasaccharide only via a single water molecule. LTB-specific residues are shown in blue, labels ended in # mark residues from a neighboring subunit in the B-pentamer. Hydrogen bonds are indicated by dashed lines (figure generated with Swiss-PdbViewer and POV-Ray). B factors for the water molecules W1–W7 are 40 Å^2 (W1), 35 Å^2 (W2), 40 Å^2 (W3), 33 Å^2 (W4), 38 Å^2 (W5), 46 Å^2 (W6), and 23 Å^2 (W7).

the saccharide ligand, thereby pulling along the two adjacent amino acid residues.

Apart from this small change in conformation, structural deviations in the binding site are mainly restricted to a few differences in sequence between LCTB*K* and the native toxins, differences that appear to affect the network of structural water molecules at this site (Figure 2B, Tables 2 and 3): while in all but three of the 25 investigated CT B-subunits, two of the central water molecules (W1 and W2) are present and form the same water-protein hydrogen bonds as in the LCTB*K* complex, the water molecule interacting with Asn4# (W1) is totally absent in all LTB structures. In contrast, none of the four water molecules (W4–W7) mediating hydrogen bonds to Fucα2 or GlcNAcβ have any resemblance in CTB, whereas they are extremely well conserved in the five high-resolution LTB structures investigated. The water molecule coordinated to the Gln3# backbone (W2) is conserved in both CTB and LTB, while the water molecule coordinated to Glu7# (W3) is characterized by weak electron density and is not well conserved in either of the parent B subunits.

Further, it was noted that a large number of CTB and LTB subunits feature water molecules (or even larger molecules such as glycerol) at positions corresponding to the hydroxyl groups and the ring oxygen of Fucα2 in the LCTB*K* complex. Similar observations have been reported for other proteins, e.g., for various legume lectin structures (Loris et al., 1994), which also exhibit conserved water molecules in carbohydrate binding sites.

Effect of S4N Mutation in hLTB

Given that only one amino acid residue (at position 4) differs between LCTB*K* and a protein hybrid with binding properties indistinguishable from hLTB (LCTB*H*), the effect of a single substitution in hLTB from Ser4 (present in hLTB and LCTB*H*) to Asn4 (found in CTB and LCTB*K*) was investigated by solid phase glycosphingolipid binding studies. The results are shown in Figure 5. Both hLTB and hLTB/S4N B-pentamers bound to the GM1 ganglioside (Galβ3GalNAcβ4(NeuAcα3)Galβ4Glcβ1Cer) and to paragloboside (Galβ4GlcNAcβ3Galβ4Glcβ1Cer) with similar affinities, demonstrating that the mutation introduced in hLTB/S4N had not affected the ability to interact with these compounds. However, with respect to the blood group antigens, the S4N mutant clearly binds much stronger to the A9 type 2 (A9-2) glycosphingolipid GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ3Galβ4GlcNAcβ3Galβ4Glcβ1Cer) than hLTB (if hLTB is at all binding). Binding to the blood group antigen is, however, weaker than to GM1. Similar results were obtained previously for LCTB*K* (Ångström et al., 2000).

Discussion

It is clear from the binding studies performed on LCTB*H* and LCTB*K* (Ångström et al., 2000) that the asparagine residue at position 4 plays a significant role for the recognition and binding of blood group A and B antigens by LCTB*K*. The Ser→Asn substitution at position 4 in the recombinant hLTB has now been shown to generate

Table 2. Ligand Contacts

Carbohydrate Residue	Direct Contacts (Å)	Interaction Partner	Indirect Contacts (Å)	Interaction Partner
GalNAcα N2	3.0 ± 0.1 (10)	Gly45 O		
GalNAcα N2	2.9 ± 0.1 (10)	Thr47 OG1		
GalNAcα O3	2.7 ± 0.1(10)	W2	2.6 ± 0.1 (9)	Gln3# O
		W2	2.8 ± 0.1 (9)	Thr47 OG1
GalNAcα O3	2.5 ± 0.2 (6)	W1	2.8 ± 0.2 (7)	Asn4# OD1/ND2
GalNAcα O4	2.7 ± 0.1 (6)	W3	2.8 ± 0.1 (6)	Glu7# OE1/NE2
Galβ4 O4	2.8 ± 0.1 (10)	Gly45 O		
Fucα2 O2	2.8 ± 0.3 (10)	Gln3# OE1/NE2		
Fucα2 O3	2.7 ± 0.1 (8)	W4	2.9 ± 0.3 (6)	Thr92 OG1
		W4	2.7 ± 0.1 (5)	W5
Fucα2 O4	2.6 ± 0.1 (10)	Thr47 O		
Fucα2 O4	2.9 ± 0.1 (10)	Asn94 N		
Fucα2 C6	3.6 ± 0.1 (10)	Phe48 CE1/CE2		
	3.9 ± 0.1 (10)	Phe48 CZ		
GlcNAcβ O1	2.9 ± 0.3 (6)	W6	3.0 ± 0.3 (5)	Gln16 OE1/NE2
		W6	3.0 ± 0.3 (5)	W7
GlcNAcβ O6	3.1[a]	W4		
GlcNAcβ O6	3.1[a]	W5	3.1 ± 0.2 (5)	Asn94 OD1/ND2
Fucα3 O4	3.2 ± 0.3 (10)[b]	Asn 44 O		
Fucα3 C6	3.9 ± 0.3 (10)	Gly 45 C		
Fucα3 C6	3.6 ± 0.1 (10)	Ala 46 CA		
	3.5 ± 0.1 (10)	Ala 46 CB		
		W7	3.1 ± 0.2 (7)	Gln16 OE1/NE2
		W7	2.7 ± 0.4 (10)	Tyr18 OH
		W7	2.6 ± 0.1 (10)	Asn94 OD1/ND2
		Asn4# OD1/ND2	2.9 ± 0.1 (10)	Glu7# N
		Asn4# OD1/ND2	3.0 ± 0.0 (9)	Thr6# N

The numbers in parentheses refer to the number of binding sites in which the particular interaction is present (from in total 10 independent sites in the asymmetric unit).
[a] Present only in subunit A.
[b] Unfavorable angle for a hydrogen bond.

a mutant (hLTB/S4N) with binding properties similar to LCTB*K* (Figure 5), giving further support to this observation. Remarkably, in the crystal structure of LCTB*K*, this residue does not interact directly with the blood group A pentasaccharide, as one might expect, but only indirectly via a single water molecule (see Figures 2B and 4). The molecular interaction involves the amide group of Asn4 and the GalNAcα3 3-OH group of the blood group A pentasaccharide. This is in contrast to the model proposed by Ångström et al. (Ångström et al., 2000), which predicted direct H-bonding interactions of Asn4 to Galβ4 and Fucα3. While the rough localization of the blood group A pentasaccharide in the vicinity of the Asn4 residue was correctly predicted, the previously proposed molecular model differs from the present experimental results in terms of overall conformation and orientation of the saccharide ligand. In this context it should be noted that, owing to the presence of ten distinct and individually refined B subunits, it is possible to exclude that the architecture of the binding site, as found in the crystal structure, is an artifact from packing and crystal contacts.

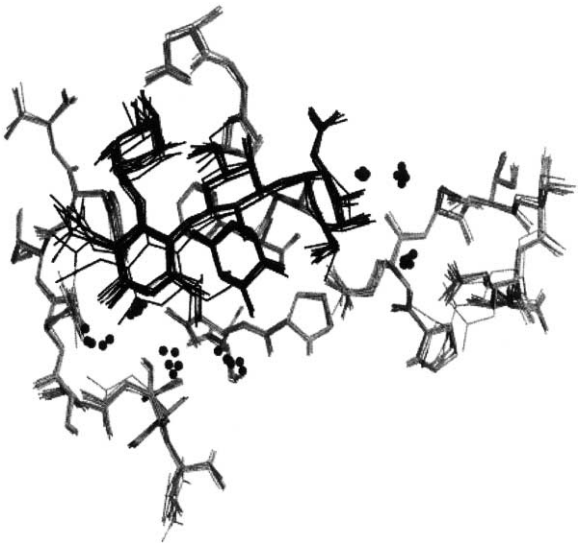

Figure 3. Superimposition of All Ten Independently Refined Blood Group A Binding Sites

The one outlier—the site at the subunit interface E/D—is strongly affected by crystal contacts. Please note the clusters of conserved water molecules bound to the site (figure generated with Swiss-PdbViewer and POV-Ray).

Role of Asn4 for Binding Blood Group Antigens

Why is Asn4 so crucial, given that it only interacts with the pentasaccharide via a single water molecule? And why can serine not fulfill the same function as asparagine in the first place? Even though the two side chains differ, they also exhibit significant similarities, such as their polar functional groups that have the potential to form water-mediated hydrogen bonds to the saccharide ligand. The clue probably lies in the double interaction of the Asn4 side chain. In all CTB and LTB structures,

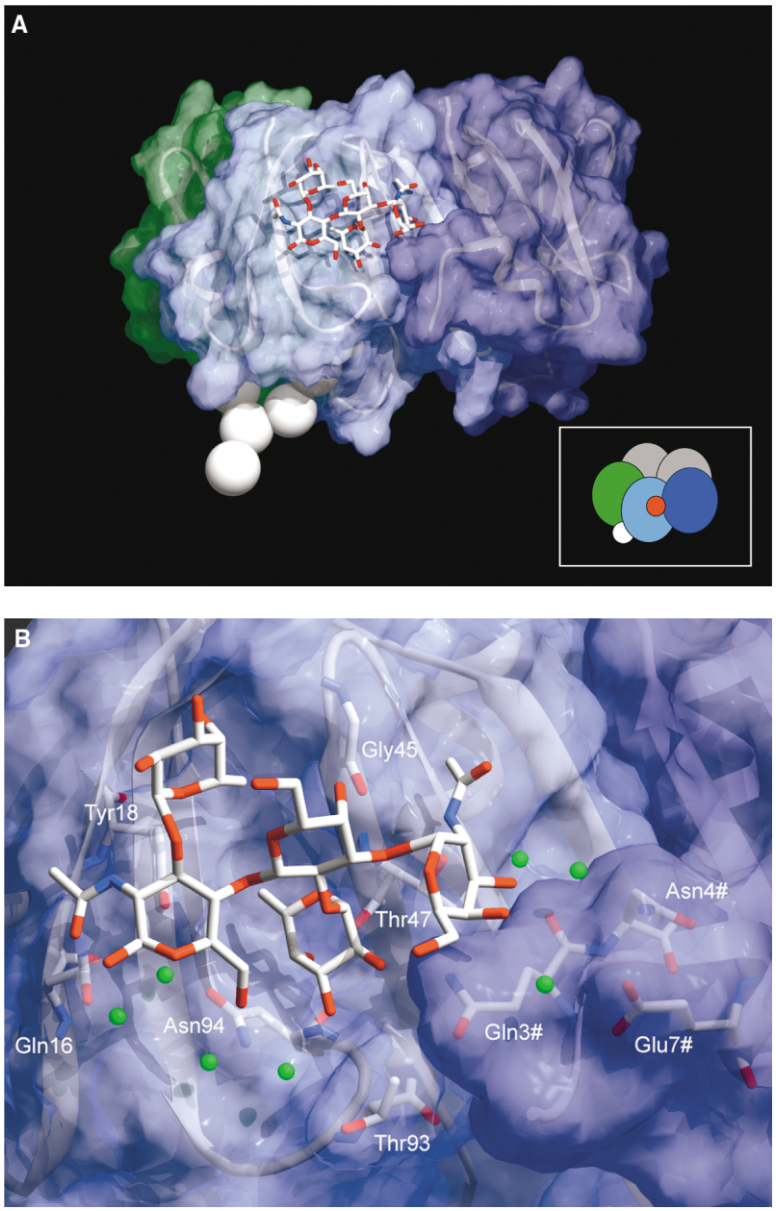

Figure 4. View of the Novel Binding Site in a Shallow Groove at the Interface of Two B Subunits of LCTB*K*

(A) Overview picture. The blood group A pentasaccharide is shown in stick representation; the GM1 pentasaccharide is indicated by white spheres. (Inset: positions of the novel binding site and the GM1 binding site within the B-pentamer, marked with red and white circles, respectively). Please note that the carbohydrate ligand is only displayed for one of the five equivalent binding sites within the B-pentamer.

(B) Close-up view of the blood group A binding site, featuring the two separate water networks (figure generated with Swiss-PdbViewer and POV-Ray).

the respective side chains are strongly coordinated to the backbone of Thr6 and Glu7, which exhibit a similar conformation in LCTB*K* compared to the parent toxins. Serine only has one functional group, a hydroxyl group. Since this group is rather tightly bound to the protein backbone of Thr6/Glu7, it is unavailable for either direct or indirect hydrogen bonding interactions with the saccharide. In contrast, Asn4, with its bifunctional amide group, can mediate interactions both to the backbone of Thr6/Glu7 and in addition to the bridging water molecule coordinated to the 3-OH group of GalNAcα3. In this way, Asn4, but not Ser4 (or Thr4, as in porcine LTB isolates) is able to provide the additional interaction required for binding of the new ligand. Molecular Dynamics Simulations (B. Källebring, personal communication) are in support of this interpretation.

While the above reasoning indicates why Asn4, but not Ser4, can contribute to ligand binding, the question remains how strong the contribution of this water-mediated interaction is in reality. More sensitive binding studies (e.g., ITC measurements) on mutant and native toxin

Table 3. Amino Acid Composition in the Novel Binding Site

Position	LCTBK/hLTB-S4N	hLTB	CTB
3	Gln	Gln	Gln
4	***Asn***	Ser[a]	***Asn***
7	***Glu***	***Glu***	Asp
16	Gln	Gln	Gln
18	***Tyr***	***Tyr***	His
45	Gly	Gly	Gly
47	Thr	Thr	Thr
92	Thr	Thr	Thr
94	***Asn***	***Asn***	His

Residues highlighted in bold italic differ between hLTB and CTB.
[a]Heat-labile enterotoxin B subunits from porcine isolates (pLTB) have a threonine residue at position 4.

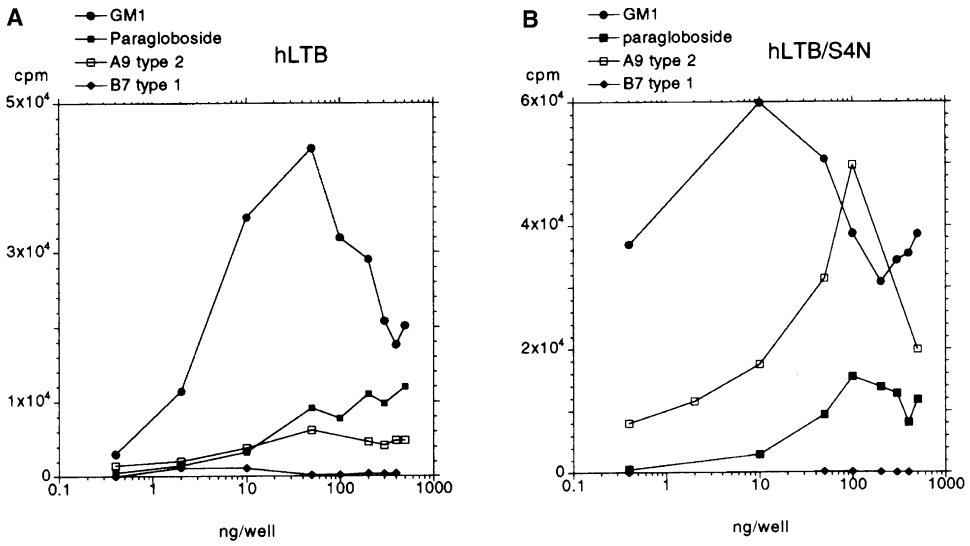

Figure 5. Glycosphingolipid Binding Data

^{125}I-labeled hLTB (A) and hLTB/S4N (B) were added to serial dilutions of glycosphingolipids adsorbed to microtiter wells. The assay was done as described in the Experimental Procedures section. Data are expressed as mean values of triplicate determinations. GM1, Galβ3GalNAcβ4 (NeuAcα3)Galβ4Glcβ1Cer; paragloboside, Galβ4GlcNAcβ3Galβ4Glcβ1Cer; A9 type 2 (A9-2), GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ3Galβ 4GlcNAcβ3Galβ4Glcβ1Cer; B7 type 1, Galα3(Fucα2)Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ1Cer.

B subunits will certainly be needed in order to fully resolve this issue. However, published data on unrelated protein-ligand complexes (Pebay-Peyroula et al., 2003; Clarke et al., 2001; Adhikari et al., 2001; Pratap et al., 2001; Ravishankar et al., 1997; Swaminathan et al., 1998) suggest that a single water-mediated interaction can in fact enhance binding affinity by up to 20-fold, which in some cases is enough for generating binding specificity (Ravishankar et al., 1997; Adhikari et al., 2001; Pratap et al., 2001; Pebay-Peyroula et al., 2003).

Importance of the Water Network

Based on the previous considerations, we estimate that the indirect, water-mediated interaction of Asn4 enhances binding affinity to the blood group A pentasaccharide approximately by one order of magnitude. Obviously, additional interactions are required to create a new binding site, especially for such large ligands as the blood group antigens. Nevertheless, provided that the essential preconditions already exist, a single substitution might be all that is needed for the creation a new binding site.

The LCTB*K*-blood group A pentasaccharide complex is characterized by very few direct interactions. More than 50% of the interactions are indirect water-mediated contacts. Still, the electron density is exceptionally clear for the complete pentasaccharide and in all ten individual copies of the binding site, indicating that the water network provides a stable platform of hydrogen bonds between the protein and its sugar ligand.

The presence of water-mediated interactions is not at all uncommon in protein-ligand recognition (Janin, 1999; Levitt and Park, 1993; Toone, 1994). Several investigations have shown that certain water molecules occupy the same position in protein structures regardless of crystallization conditions and for proteins from different species, provided that the proteins are structurally related (Ogata and Wodak, 2002; Loris et al., 1994; Krem and Di Cera, 1998; Shaltiel et al., 1998; Sreenivasan and Axelsson, 1992). Furthermore, these conserved water molecules are often found in active site clefts (Shaltiel et al., 1998) or in ligand binding sites (Poornima and Dean, 1995a, 1995b, 1995c; Babor et al., 2002) and can be considered to be an extension of the protein surface.

As pointed out in the Results section, the blood group A binding site of LCTB*K* harbors two distinct water networks (Figures 2B, 3, and 4). One of them (W1-W3) mediates interactions between the terminal carbohydrate moiety GalNAcα3 and the adjacent B subunit—with Asn4 at the center of attention. The second network (W4-W7) consists of four water molecules that primarily mediate interactions to the reducing end of the saccharide, GlcNAc, but also to Fucα2. When compared with the native toxin B subunits, LTB and CTB, a clear sequence dependence of this water coordination was noted. The water molecule coordinated by Asn4 in the crystal structure (W1) was totally absent in all LTB structures, while well conserved in CTB. On the contrary, the water network coordinated by residues Gln16, Tyr18, Thr92, and Asn94 (W4-W7) is well preserved in LTB but has no similarities in CTB. This correlates well with the sequence differences in LTB and CTB (see Table 3, Figure 2B). The only CTB-specific side chain participating in any kind of interaction with the sugar is Asn4. All other residues at this site are either hLTB specific or conserved between the two toxins. The blood group A binding site of LCTB*K* thus corresponds exactly to that obtained by the introduction of a single amino acid substitution (S4N) in hLTB, which resulted in a molecule with similar binding properties.

Function of Second Binding Site?

The indications are that the novel ligand binding site was essentially preformed by the large water network

held in place by the LTB residues and that the single substitution of Ser4 by asparagine was all that was needed to strengthen the binding affinity such that it became detectable by our assays. More sensitive binding studies will be needed to investigate this question further. If experimentally validated, this speculation leads to further questions regarding the evolution of the binding site. Whereas it may be unlikely that an essentially preformed nonfunctional binding site exists awaiting a single mutation to be activated, it is probable that this second binding site already has some function, possibly partly explaining the broader carbohydrate specificity observed for LTB. Interestingly, a naturally occurring *E. coli* heat-labile enterotoxin B subunit homolog with a 100% sequence identity to LCTBK in the described new binding site was recently described (Karasawa et al., 2002). This finding needs to be followed up by sequencing the B subunits from a large number of enterotoxigenic *E.coli* to determine the frequency of Asn4.

There is ample evidence for the ability of LTB to bind to additional *N*-acetyllactosamine-based receptors via uncharacterized mechanisms that do not necessarily involve the GM1 binding site (Griffiths and Critchley, 1991; Holmgren et al., 1982, 1985; Karlsson et al., 1996; Shida et al., 1996; Teneberg et al., 2000). Indications that LTB is able to bind to blood group A/B determinants of both glycolipid (Barra et al., 1992) and glycoprotein (Balanzino et al., 1994) origin have also been described. Furthermore, one of these studies (Holmgren et al., 1982) shows that LT remains toxic even if the primary GM1 binding sites are blocked. Since five of the eight nonconserved residues exchanged in the block substitution generating the CTB mutant LCTB*H* (Bäckström et al., 1997) (with binding properties indistinguishable from hLTB) are found in the novel binding site and four of these residues interact with the new ligand, it is reasonable to assume that the broader range of specificities observed for LTB at least in part relates to the newly discovered binding site. Interestingly, the LT toxin, which remains associated with the outer membrane lipopolysaccharides (LPS) of enterotoxigenic *E. coli* prior to attaching to the host cell, has been suggested to exhibit an LPS binding site in the same general region as the novel blood group A/B binding site (Horstman et al., 2004).

In this context, the question whether the newly discovered binding site is membrane accessible or not merits consideration, i.e., if blood group antigens displayed on the cell surface at all could reach this binding site. This is highly relevant since the native heat-labile enterotoxin has been shown to bind to both lipid- and protein-anchored blood group antigens originating from porcine intestinal mucosa and HT-29 colon carcinoma cells (Barra et al., 1992; Balanzino et al., 1994; Galván et al., 2004). Simple docking studies (data not shown) reveal that longer blood group antigens such as the A9 type 2 glycosphingolipid may very well assume conformations extending from the membrane into the described binding site without clashing with the GM1 binding sites. However, in the human target tissue, i.e., the small intestinal epithelium, the absolute majority of nonacid blood group determinant-carrying glycosphingolipids are based on type 1 chains. Blood group determinants based on type 2 chains are, on the other hand, abundant on glycoproteins (Finne et al., 1989). For this reason, one might speculate that glycoprotein binding to the described binding site may occur either as a first step guiding the toxin to the primary GM1 sites on the epithelial cell surface or as a second step enhancing the primary binding. Alternatively, binding to this site might be totally self-sufficient for intoxication. In either scenario, the A/B determinant binding sites could, for example, play a role for the internalization process of the toxin through rafts.

Binding of Different Blood Group Determinants

According to previous binding studies (Ångström et al., 2000), LCTB*K* recognizes both the blood group A (**GalNAc**α3(Fucα2)Galβ-) and B (**Gal**α3(Fucα2)Galβ-) determinants on type 2 core chains, i.e., the minimum binding epitope required for binding of LCTB*K* is Galα3(Fucα2)Galβ4(Fucα3)GlcNAcβ (see Figure 1). Blood group A determinants exhibit an extra acetamido group attached to Galα3 at the 2′-position, which, according to the results from the binding studies is not essential for the protein-carbohydrate interactions. In the crystal structure, the acetamido nitrogen is involved in strong hydrogen bonds to two amino acid residues, Gly45 and Thr47. However, the hydroxyl group present in blood group B determinants preserves this interaction, hence explaining the similar binding results.

While A9 and A7 type 2 antigens were tested binding positive, A6 (or B6) type 2 antigens were found not to bind to LCTB*K* (Ångström et al., 2000). A6 and B6 type 2 antigens also carry blood group A and B determinants, but lack the internal **Fucα3**. According to the crystal structure of the LCTB*K* complex, this α3-linked fucose is highly solvent exposed and not involved in any hydrogen bonding interactions. However, the fucose methyl group is instead involved in strong hydrophobic interactions with an extended hydrophobic surface stretch on the peptide chain, extending all the way from the Ala46 methyl group to its main chain α carbon atom and further to Gly45 C/CA (for details, see Table 2). In addition, Fucα3 imposes rigidity on the adjacent β1-4 linkage through internal carbohydrate-carbohydrate interactions with Gal and GlcNAc. The internal Fucα3 thus contributes to binding in a dual fashion: first, through direct hydrophobic interactions with the protein and, second, by internally stabilizing the conformation of the GlcNAc residue, thus priming it for binding.

Finally, the question arises as to why only blood group determinants on type 2 core chains, in which the galactose and fucose extensions at C3 and C4 of GlcNAc are interchanged relative to type 1 core chains, are recognized. Superimposing the type 1 and type 2 blood group A pentasaccharides (GalNAcα3(Fucα2)Gal**β3/4**(Fuc**α4/3**)GlcNAcβ) (Figure 1) reveals that only the GlcNAc residue suffers a 180° reorientation in the type 1 case relative to the type 2 case (cf. Teneberg et al., 2003), while the terminal trisaccharide is superimposable. The Lewis fucose (Fucα3 in type 2 and Fucα4 in type 1 chains) also occupies a similar position in both antigens. Docking studies show that even the placement of the GlcNAc *N*-acetyl group in type 1 chains at the position of the

hydroxy methyl group in type 2 chains does not seem critical, as it merely would displace one water molecule and reorient a few other ones in the binding site, although the quantitative effect of this difference is difficult to judge. Alternatively, the critical difference might reside in the positioning of the GlcNAc O1 atom, which determines the direction of the core extension of the saccharide chain. For A7, having a core extension of 3Galβ4Glcβ1Cer instead of 3Galβ4GlcNAcβ3Galβ4Glcβ1Cer as for A9, the type 1 variant may produce some minor clashes with the protein depending on the preferred GlcNAcβ3Gal conformation. This is in contrast to the weakly binding type 2 variant, which might be sufficient to render this structure nonbinding. However, for the A9 type 1 determinant, irrespective of whether the core tetrasaccharide is type 1 or type 2, binding could possibly occur as judged by docking of these structures. Binding studies with A9 type 1 variants have so far not been performed but would be interesting to test.

Implications for Drug Design

The structure of the LCTB*K*-blood group A pentasaccharide complex has brought new insights into the complex nature of protein-ligand interactions, in particular revealing the importance of water-mediated interactions for protein-carbohydrate complexes. This could prove to be important for the theoretical design of new binding specificities in lectins in general but also for any type of structure-based drug design. With regard to the creation of new recognition sites, CTB- and LTB-related proteins are, due to their extensively described structures and exceptional immunological properties, excellent targets for investigations aimed at giving insights into how their diverse biological activities are related to recognition of different carbohydrate receptors. In terms of future use of LT-related holotoxins or lectin moieties as immune modulators in disease, it is also important to further characterize the biological effect of new binding specificities and to evaluate the importance of binding to different carbohydrate receptors to the biological action of the protein.

Conclusions

The driving force for the current crystallographic analysis was the discovery of a novel binding specificity of the protein hybrid LCTB*K* to blood group antigens, based on glycosphingolipid binding studies. The 1.9 Å structure of LCTB*K* in complex with the blood group A type 2 pentasaccharide reveals the existence of a distinct new binding site, as previously predicted by molecular modeling. In the present study, we show that this new binding site is the effect of a single Ser→Asn substitution in hLTB. Unexpectedly, the residue crucial for creating the novel binding specificity, Asn4, plays only a modest role in ligand binding, interacting with the pentasaccharide ligand via a single water molecule. Also, most of the other ligand contacts are of an indirect nature, mediated by strongly bound water molecules, confirming the importance of water-mediated bonds for carbohydrate-protein interactions.

There is ample evidence in the literature indicating that LT possibly exhibits additional binding site(s) unrelated to the primary GM1 receptor site. As discussed in detail above, it might well turn out that the newly discovered binding site is in fact functional and explains several of these previously puzzling observations.

Experimental Procedures

Plasmids and DNA Manipulations

The construction of the gene encoding LCTB*K* has been described previously (Ångström et al., 2000). The gene encoding human LTB with a single amino acid change at position 4 (S4N) was derived from the plasmid encoding the LCTB*K* protein. The plasmid pML-LCTB*K* was digested with PstI and HindIII to remove the entire LTB/CTB hybrid downstream of the cysteine at position 9 of the mature protein. DNA encoding the native human LTB gene, obtained by PCR amplification of the *heltB* gene, was digested with the same enzymes and replaced the DNA downstream of position 9. Following ligation of the DNA fragments and their transformation into the classical O1 *Vibrio cholerae* strain JS1569, recombinant plasmids were isolated on the basis first of restriction analysis and finally by confirmation of the DNA sequence using an A310 Genetic analyzer in combination with Big Dye sequencing chemistry (Perkin Elmer). The recombinant hLTB/S4N could then be expressed and purified essentially in the same way as LCTB*K*.

Production, Purification, and Characterization of the Hybrid and Mutant B Subunits

The plasmids encoding the mutant and hybrid B subunits were expressed in *Vibrio cholerae* strain JS1569. Cell cultures were grown for 20–24 hr at 37°C in liquid cultures of modified Syncase medium supplemented with 100 μg/ml ampicillin. The bacteria were removed by centrifugation and the B subunits, which are secreted into the growth medium, were recovered from the supernatant by acidification in the presence of 2.5 g/l sodium hexametaphosphate (Lebens et al., 1993). The precipitate was redissolved in 20–200 mM Tris/HCl, pH 7.5–8.0. The protein was further purified using either gel filtration or ion exchange chromatography, depending on the field of application. For microtiter well assays, LCTB*K* and hLTB/S4N were purified on a Superdex 200 gel filtration column in 10 mM Tris/HCl, pH 7.5, and concentrated to a final concentration of 1 mg/ml. Since the protein sample prepared by this procedure proved to be of insufficient quality for crystallization trials, protein for crystallization was instead purified using a BioRex weak cationic ion exchange resin in 20 mM Tris/HCl, pH 8.0, and eluted using a salt gradient of 0–0.5 M NaCl in the same buffer. The protein was subsequently dialyzed against Tris/HCl, pH 7.5, supplemented with 0.2 M NaCl, and concentrated to a final concentration of 9 mg/ml.

The concentration of recombinant B subunits was monitored by GM1 enzyme-linked immunosorbent assays (GM1-ELISA) using the crossreactive monoclonal antibody LT39 and recombinant CTB of a known concentration as a standard (Svennerholm and Holmgren, 1978). The protein was further analyzed by SDS-polyacrylamide gel electrophoresis and the concentration of pure protein was determined with the BCA Protein Assay (PIERCE) and by absorbance spectroscopy at 280 nm.

Crystallization

The LCTB*K* hybrid (9 mg/ml in Tris/HCl, pH 7.5, 0.2 M NaCl) was cocrystallized with the terminal pentasaccharide of the A9 type 2 (A9-2) glycosphingolipid GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ (IsoSep AB; Tullinge, Sweden). Crystals were obtained from freshly prepared protein sample, using the hanging drop vapor diffusion technique at room temperature. Saccharide and protein B subunits were mixed in a molar ratio of approximately 3:1 and the solution was allowed to equilibrate for ca. 2 hr prior to set-ups. The reservoir solution contained 0.25–0.30 M $CaCl_2$, 23%–24% PEG 3350, and 20% glycerol. The crystals belong to the monoclinic space group C2, with cell axes a = 110.2 Å, b = 70.1 Å, c = 137.4 Å and β = 92.9°. A Matthews coefficient V_M of 2.4 $Å^3$/Da (Matthews, 1968) corresponding to a solvent content of about 48.5% was obtained for two

copies of the B-pentamer in the asymmetric unit, i.e., ten copies of the B subunit protomers.

Data Collection

Data sets were collected for two different crystals at cryogenic temperature (100 K) and used in different stages of structure determination and refinement. One of the data sets was recorded to 2.0 Å resolution on a MAR CCD detector at beamline BL711 at the MAX-lab II synchrotron in Lund, Sweden (400 frames; oscillation range of 0.5° per frame; wavelength $\lambda = 0.97$ Å). The second data set was recorded to 1.9 Å resolution on a MAR 345 imaging-plate system mounted on a rotating copper anode (Rigaku RU300 HB) at AstraZeneca, Mölndal (300 frames; oscillation range of 1° per frame).

The data were processed and scaled using XDS and XSCALE (Kabsch, 1988a, 1988b, 1993). The former data set was used for structure determination and initial refinement. The 1.9 Å rotating anode data were used for subsequent refinement owing to the slightly higher resolution, higher signal-to-noise ratio, and lower final R factors. This data set was also used for calculation of the final model. Attempts were made to scale the two data sets together, but without success. For statistics of both data sets, see Table 1.

Structure Determination and Refinement

The structure was solved by Molecular Replacement using the program AMoRe (Navaza, 1994), as implemented in the CCP4 program suite (CCP4, 1994). The cholera toxin B-pentamer from PDB entry 1EEI (Fan et al., 2001) was used as a search model. A self-rotation search with the CCP4 program MOLREP (Vagin and Teplyakov, 1997) revealed a 2-fold noncrystallographic symmetry axis in addition to the expected 5-fold noncrystallographic symmetry, consistent with the presence of *two* pentamers per asymmetric unit. Two molecular replacement solutions were found for the pentamer search model, and the resulting decamer was used as initial model for refinement. Rigid body optimization resulted in unbiased, but clearly interpretable electron density for the five sugar residues in all ten copies of the B subunit, already straight from the start. Crystallographic refinement using the program CNS (Brünger et al., 1998) was then alternated with manual rebuilding of the coordinates using the program O (Jones et al., 1991), in several cycles. Initially, the 10-fold noncrystallographic symmetry (NCS) was used to constrain and, subsequently, to restrain the subunits to be similar. To aid rebuilding in initial stages, an NCS-averaged map was calculated with CNS, and no more than the features clearly visible in this map were introduced to the monomer model. Later, the restraints were loosened and each monomer and its residues were refined individually. Throughout refinement, the R_{free} value, based on 5% of the collected data, was used to monitor the convergence.

Despite all efforts to improve the model to fit the experimental data, R and R_{free} never decreased below 30% and 32%, respectively, until we turned to the data from the home source for further refinement. The incomplete LCTB*K* model, from which the oligosaccharides had been removed, was taken as a starting model for refinement. After rigid body minimization and one step each of NCS-phased refinement (CCP4) and restrained refinement using the CCP4 supported program Refmac5 (Murshudov et al., 1997), the R factors had decreased significantly. Subsequently, the carbohydrate moieties were added back to the model (using the coordinates from the modeling studies [Ångström et al., 2000]). The parameter and topology file for GlcNAcβ was generated using the CCP4 supported program LIBCHECK, while the files for the other saccharide units were already present in the monomer library. In later stages, the ARP_water feature in Refmac5 and ARP/wARP was used for building the solvent atoms. Criteria for adding water molecules were, in addition to forming reasonable hydrogen bonds, strong peaks in the $F_o - F_c$ map ($>3.5\sigma$). All manual rebuilding was done with O, using composite simulated annealed OMIT maps calculated using CNS for validation, in addition to the ordinary $2F_o - F_c$ and $F_o - F_c$ electron density maps. Refinement statistics are summarized in Table 1.

Structural Comparisons

Structural comparisons were performed using the CCP4 program LSQMAN (Kleywegt, 1996). Each of the ten B subunits in the crystal structure was consecutively superimposed on the others and the root-mean-square deviation (rmsd) was calculated for the backbone and for each of the ten pentasaccharides. Comparisons with the native toxins were based on superimpositions with high-resolution structures of LTB and CTB (native and mutant) deposited in the Protein Data Bank (PDB). Ligand-bound complexes as well as genetically modified molecules were included, provided that there were no relevant changes in the sequence and the resolution was sufficiently high (≤2 Å). In total, ten structures were selected from the PDB for this comparison (ID codes 1DJR, 1EEF, 1EFI, 1FD7, 1LT5, 3CHB, 1LLR, 1MD2, 1JR0, and 1EEI). From these, five CTB structures (Fan et al., 2001; Merritt et al., 1998, 2002; Pickens et al., 2002; Zhang et al., 2002) and five LTB structures (Fan et al., 2001; Merritt et al., 1997; Minke et al., 2000) were investigated further to analyze conserved water molecules.

Reference Glycosphingolipids

Total acid and nonacid glycosphingolipid fractions were obtained by standard procedures (Karlsson and Stromberg, 1987). The individual glycosphingolipids were isolated by repeated chromatography on silicic acid columns and by HPLC of the native glycosphingolipid fractions, or acetylated derivatives thereof. The identity of the purified glycosphingolipids was confirmed by mass spectrometry (Samuelsson et al., 1990), proton NMR spectroscopy (Koerner et al., 1983), and degradation studies (Yang and Hakomori, 1971).

Radiolabeling

Aliquots of 100 μg of each B subunit preparation were labeled with ^{125}I by the Iodogen method (Laemmli, 1970), giving in average 2×10^3 cpm/μg.

Microtiter Well Assay

The microtiter well binding assay was performed as previously described (Ångström et al., 1994). In short, 50 μl of serial dilutions (each dilution in triplicate) of pure glycosphingolipids in methanol were applied to microtiter wells (Falcon 3911, Becton Dickinson Labware, Oxnard, CA). When the solvent had evaporated, the wells were blocked for 2 hr at room temperature with 200 μl of phosphate-buffered saline, pH 7.3, (PBS) containing 2% bovine serum albumin (w/v) and 0.1% NaN_3 (w/v) (Sol. 1). Thereafter, 50 μl of radiolabeled B subunits, diluted in Sol. 1 (approximately 2×10^3 cpm/μl), were added per well and incubated over night at room temperature. After washing six times with PBS, the wells were cut out and the radioactivity was counted in a γ counter.

Acknowledgments

We wish to thank Hanna Jansson for spending a lot of time helping with crystallization set-ups in the starting phase of the project and Cecilia Cronet for help and support throughout the protein preparation process and initial refinement of the crystal structure. Bruno Källebring is gratefully acknowledged for performing Molecular Dynamics Simulations on this system. We would further like to thank Helena Käck and Tomas Lundqvist for data collection time and assistance at AstraZeneca and Yngve Cerenius for support at MAX-lab in Lund. This study has been supported by grants from the Glycoconjugates in Biological Systems program from the Swedish National Foundation for Strategic Research (research position of U.K.), Hagbergs Foundation (U.K.), the Swedish Research Council (grants no. 621-2003-4057 [U.K.]; 12628 [S.T.]; 97-296 [M.L.]), the Swedish Cancer Foundation, the Swedish Medical Society, and the Chalmers Bioscience Initiative.

Received: February 8, 2004
Revised: June 10, 2004
Accepted: June 24, 2004
Published: September 7, 2004

References

Adhikari, P., Bachhawat-Sikder, K., Thomas, C.J., Ravishankar, R., Jeyaprakash, A.A., Sharma, V., Vijayan, M., and Surolia, A. (2001). Mutational analysis at Asn-41 in peanut agglutinin. A residue critical

for the binding of the tumor-associated Thomsen-Friedenreich antigen. J. Biol. Chem. *276*, 40734–40739.

Alper, J. (2001). Searching for medicine's sweet spot. Science *291*, 2338–2343.

Ångström, J., Teneberg, S., and Karlsson, K.A. (1994). Delineation and comparison of ganglioside-binding epitopes for the toxins of *Vibrio cholerae*, Escherichia coli, and Clostridium tetani: evidence for overlapping epitopes. Proc. Natl. Acad. Sci. USA *91*, 11859–11863.

Ångström, J., Bäckström, M., Berntsson, A., Karlsson, N., Holmgren, J., Karlsson, K.A., Lebens, M., and Teneberg, S. (2000). Novel carbohydrate binding site recognizing blood group A and B determinants in a hybrid of cholera toxin and *Escherichia coli* heat-labile enterotoxin B-subunits. J. Biol. Chem. *275*, 3231–3238.

Babor, M., Sobolev, V., and Edelman, M. (2002). Conserved positions for ribose recognition: importance of water bridging interactions among ATP, ADP and FAD-protein complexes. J. Mol. Biol. *323*, 523–532.

Bäckström, M., Shahabi, V., Johansson, S., Teneberg, S., Kjellberg, A., Miller-Podraza, H., Holmgren, J., and Lebens, M. (1997). Structural basis for differential receptor binding of cholera and *Escherichia coli* heat-labile toxins: influence of heterologous amino acid substitutions in the cholera B-subunit. Mol. Microbiol. *24*, 489–497.

Balanzino, L.E., Barra, J.L., Monferran, C.G., and Cumar, F.A. (1994). Differential interaction of *Escherichia coli* heat-labile toxin and cholera toxin with pig intestinal brush border glycoproteins depending on their ABH and related blood group antigenic determinants. Infect. Immun. *62*, 1460–1464.

Barra, J.L., Monferran, C.G., Balanzino, L.E., and Cumar, F.A. (1992). *Escherichia coli* enterotoxin preferentially interacts with blood group A-active glycolipids from pig intestinal mucosa and A- and B-active glycolipids from human red cells compared to H-active glycolipids. Mol. Cell. Biochem. *115*, 63–70.

Beachey, E.H. (1981). Bacterial adherence: adhesin-receptor interactions mediating the attachment of bacteria to mucosal surface. J. Infect. Dis. *143*, 325–345.

Belisle, B.W., Twiddy, E.M., and Holmes, R.K. (1984). Monoclonal antibodies with an expanded repertoire of specificities and potent neutralizing activity for *Escherichia coli* heat-labile enterotoxin. Infect. Immun. *46*, 759–764.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D. Biol. Crystallogr. *54*, 905–921.

CCP4 (Collaborative Computational Project, Number 4). (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Clarke, C., Woods, R.J., Gluska, J., Cooper, A., Nutley, M.A., and Boons, G.J. (2001). Involvement of water in carbohydrate-protein binding. J. Am. Chem. Soc. *123*, 12238–12247.

Dennis, C. (2003). Sweet revenge. Nature *423*, 580–582.

Dove, A. (2001). The bittersweet promise of glycobiology. Nat. Biotechnol. *19*, 913–917.

Fan, E., Merritt, E.A., Zhang, Z., Pickens, J.C., Roach, C., Ahn, M., and Hol, W.G. (2001). Exploration of the GM1 receptor-binding site of heat-labile enterotoxin and cholera toxin by phenyl-ring-containing galactose derivatives. Acta Crystallogr. D Biol. Crystallogr. *57*, 201–212.

Finne, J., Breimer, M.E., Hansson, G.C., Karlsson, K.-A., Leffler, H., Vliegenthart, J.F.G., and van Halbeck, H. (1989). Novel polyfucosylated *N*-linked glycopeptides with blood group A, H, X, and Y determinants from human small intestinal epithelial cells. J. Biol. Chem. *264*, 5720–5735.

Galván, E.M., Diema, C.D., Roth, G.A., and Monferran, C.G. (2004). Ability of blood group A-active glycosphingolipids to act as *Escherichia coli* heat-labile enterotoxin receptors in HT-29 cells. J. Infect. Dis. *189*, 1556–1564.

Griffiths, S.L., and Critchley, D.R. (1991). Characterisation of the binding sites for *Escherichia coli* heat-labile toxin type I in intestinal brush borders. Biochim. Biophys. Acta *1075*, 154–161.

Gyles, C.L. (1992). *Escherichia coli* cytotoxins and enterotoxins. Can. J. Microbiol. *38*, 734–746.

Holmes, R.K., and Twiddy, E.M. (1983). Characterization of monoclonal antibodies that react with unique and cross-reacting determinants of cholera enterotoxin and its subunits. Infect. Immun. *42*, 914–923.

Holmgren, J. (1973). Comparison of the tissue receptors for *Vibrio cholerae* and *Escherichia coli* enterotoxins by means of gangliosides and natural cholera toxoid. Infect. Immun. *8*, 851–859.

Holmgren, J. (1981). Actions of cholera toxin and the prevention and treatment of cholera. Nature *292*, 413–417.

Holmgren, J., Lönnroth, I., Månsson, J., and Svennerholm, L. (1975). Interaction of cholera toxin and membrane GM1 ganglioside of small intestine. Proc. Natl. Acad. Sci. USA *72*, 2520–2524.

Holmgren, J., Fredman, P., Lindblad, M., Svennerholm, A.M., and Svennerholm, L. (1982). Rabbit intestinal glycoprotein receptor for *Escherichia coli* heat-labile enterotoxin lacking affinity for cholera toxin. Infect. Immun. *38*, 424–433.

Holmgren, J., Lindblad, M., Fredman, P., Svennerholm, L., and Myrvold, H. (1985). Comparison of receptors for cholera and *Escherichia coli* enterotoxins in human intestine. Gastroenterology *89*, 27–35.

Horstman, A.L., Bauman, S.J., and Kuehn, M.J. (2004). Lipopolysaccharide 3-deoxy-D-manno-octulosonic acid (Kdo) core determines bacterial association of secreted toxins. J. Biol. Chem. *279*, 8070–8075.

Janin, J. (1999). Wet and dry interfaces: the role of solvent in protein-protein and protein-DNA recognition. Structure *7*, R277–R279.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kabsch, W. (1988a) Automatic indexing of rotation diffraction patterns. J. Appl. Crystallogr. *21*, 67–71.

Kabsch, W. (1988b) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. J. Appl. Crystallogr. *21*, 916–924.

Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. *26*, 795–800.

Karasawa, T., Hideaki, I., Tsukamoto, T., Yamasaki, S., Kurazono, H., Faruque, S.M., Nair, G.B., Nishibuchi, M., and Takeda, Y. (2002). Cloning and characterization of genes encoding homologues of the B subunit of cholera toxin and the *Escherichia coli* heat-labile enterotoxin from clinical isolates of *Citrobacter freundii* and *E. coli*. Infect. Immun. *70*, 7153–7155.

Karlsson, K.A. (1989). Animal glycosphingolipids as membrane attachment sites for bacteria. Annu. Rev. Biochem. *58*, 309–350.

Karlsson, K.A., and Stromberg, N. (1987). Overlay and solid-phase analysis of glycolipid receptors for bacteria and viruses. Methods Enzymol. *138*, 220–232.

Karlsson, K.A., Teneberg, S., Ångström, J., Kjellberg, A., Hirst, T.R., Bergström, J., and Miller-Podraza, H. (1996). Unexpected carbohydrate cross-binding by *Escherichia coli* heat-labile enterotoxin. Recognition of human and rabbit target cell glycoconjugates in comparison with cholera toxin. Bioorg. Med. Chem. *4*, 1919–1928.

Kleywegt, G.J. (1996). Use of non-crystallographic symmetry in protein structure refinement. Acta Crystallogr. D Biol. Crystallogr. *52*, 842–857.

Koerner, T.A., Jr., Prestegard, J.H., Demou, P.C., and Yu, R.K. (1983). High-resolution proton NMR studies of gangliosides. 1. Use of homonuclear two-dimensional spin-echo J-correlated spectroscopy for determination of residue composition and anomeric configurations. Biochemistry *22*, 2676–2687.

Krem, M.M., and Di Cera, E. (1998). Conserved water molecules in the specificity pocket of serine proteases and the molecular mechanism of Na^+ binding. Proteins *30*, 34–42.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680–685.

Lai, C.-Y. (1980). The chemistry and biology of cholera toxin. CRC Crit. Rev. Biochem. *12*, 171–206.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lebens, M., Johansson, S., Osek, J., Lindblad, M., and Holmgren, J. (1993). Large-scale production of *Vibrio cholerae* toxin B subunit for use in oral vaccines. Biotechnology (N.Y.) *11*, 1574–1578.

Lebens, M., Sun, J.B., Sadeghi, H., Bäckström, M., Olsson, I., Mielcarek, N., Li, B.L., Capron, A., Czerkinsky, C., and Holmgren, J. (2003). A mucosally administered recombinant fusion protein vaccine against schistosomiasis protecting against immunopathology and infection. Vaccine *21*, 514–520.

Levitt, M., and Park, B.H. (1993). Water: now you see it, now you don't. Structure *15*, 223–226.

Loris, R., Stas, P.P.G., and Wyns, L. (1994). Conserved waters in legume lectin crystal structures. J. Biol. Chem. *269*, 26722–26733.

Maeder, T. (2002). Sweet medicine. Sci. Am. *7*, 24–31.

Matthews, B.W. (1968). Solvent content of protein crystals. J. Mol. Biol. *33*, 491–497.

Merritt, E.A., and Hol, W.G.J. (1995). AB_5 toxins. Curr. Opin. Struct. Biol. *5*, 165–171.

Merritt, E.A., Sarfaty, S., van den Akker, F., L'Hoir, C., Martial, J.A., and Hol, W.G. (1994). Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. Protein Sci. *3*, 166–175.

Merritt, E.A., Sarfaty, S., Feil, I.K., and Hol, W.G. (1997). Structural foundation for the design of receptor antagonists targeting *Escherichia coli* heat-labile enterotoxin. Structure *5*, 1485–1499.

Merritt, E.A., Kuhn, P., Sarfaty, S., Erbe, J.L., Holmes, R.K., and Hol, W.G. (1998). The 1.25 Å resolution refinement of the cholera toxin B-pentamer: evidence of peptide backbone strain at the receptor-binding site. J. Mol. Biol. *282*, 1043–1059.

Merritt, E.A., Zhang, Z., Pickens, J.C., Ahn, M., Hol, W.G., and Fan, E. (2002). Characterization and crystal structure of a high-affinity pentavalent receptor-binding inhibitor for cholera toxin and E. coli heat-labile enterotoxin. J. Am. Chem. Soc. *124*, 8818–8824.

Minke, W.E., Pickens, J., Merritt, E.A., Fan, E., Verlinde, C.L., and Hol, W.G. (2000). Structure of m-carboxyphenyl-alpha-D-galactopyranoside complexed to heat-labile enterotoxin at 1.3 Å resolution: surprising variations in ligand-binding modes. Acta Crystallogr. D Biol. Crystallogr. *56*, 795–804.

Mirelman, D., and Ofek, I. (1986). Introduction to microbial lectins and agglutinins. In Microbial Lectins and Agglutinins, D. Mirelman, ed. (New York: John Wiley & Sons), pp. 1–19.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Nashar, T.O., Williams, N.A., Hirst, T.R., and Nahar, T.O. (1996). Cross-linking of cell surface ganglioside GM1 induces the selective apoptosis of mature CD8+ T lymphocytes. Int. Immunol. *8*, 731–736.

Navaza, J. (1994). AMoRE: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Ogata, K., and Wodak, S.J. (2002). Conserved water molecules in MHC class-I molecules and their putative structural and functional roles. Protein Eng. *15*, 697–705.

Orlandi, P.A., Critchley, D.R., and Fishman, P.H. (1994). The heat-labile enterotoxin of *Escherichia coli* binds to polylactosaminoglycan-containing receptors in CaCo-2 human intestinal epithelial cells. Biochemistry *33*, 12886–12895.

Pebay-Peyroula, E., Dahout-Gonzalez, C., Kahn, R., Trezeguet, V., Lauquin, G.J.M., and Brandolin, R. (2003). Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. Nature *426*, 39–44.

Pickens, J.C., Merritt, E.A., Ahn, M., Verlinde, C.L., Hol, W.G., and Fan, E. (2002). Anchor-based design of improved cholera toxin and E. coli heat-labile enterotoxin receptor binding antagonists that display multiple binding modes. Chem. Biol. *9*, 215–224.

Poornima, C.S., and Dean, P.M. (1995a). Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. J. Comput. Aided Mol. Des. *9*, 500–512.

Poornima, C.S., and Dean, P.M. (1995b). Hydration in drug design. 2. Influence of local site surface shape on water binding. J. Comput. Aided Mol. Des. *9*, 513–520.

Poornima, C.S., and Dean, P.M. (1995c). Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. J. Comput. Aided Mol. Des. *9*, 521–531.

Pratap, J.V., Bradbrook, G.M., Reddy, G.B., Surolia, A., Raftery, J., Helliwell, J.R., and Vijayan, M. (2001). The combination of molecular dynamics with crystallography for elucidating protein-ligand interactions: a case study involving peanut lectin complexes with T-antigen and lactose. Acta Crystallogr. D Biol. Crystallogr. *57*, 1584–1594.

Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A, and Vijayan, M. (1997). Crystal structure of the peanut lectin-T-antigen complex. Carbohydrate specificity generated by water bridges. Curr. Sci. *72*, 855–861.

Sadeghi, H., Bregenholt, S., Wegmann, D., Petersen, J.S., Holmgren, J., and Lebens, M. (2002). Genetic fusion of human insulin B-chain to the B-subunit of cholera toxin enhances in vitro antigen presentation and induction of bystander suppression in vivo. Immunology *106*, 237–245.

Samuelsson, B.E., Pimlott, W., and Karlsson, K.A. (1990). Mass spectrometry of mixtures of intact glycosphingolipids. Methods Enzymol. *193*, 623–646.

Science. (2001). Review series. Science *291*, 2337–2378.

Shaltiel, S., Cox, S., and Taylor, S.S. (1998). Conserved water molecules contribute to the extensive network of interactions at the active site of protein kinase A. Proc. Natl. Acad. Sci. USA *95*, 484–491.

Sharon, N., and Lis, H. (1993). Carbohydrates in cell recognition. Sci. Am. *268*, 82–89.

Shida, K., Takamizawa, K., Takeda, T., and Osawa, T. (1996). Characterization by Western blotting of mouse intestinal glycoproteins bound by *Escherichia coli* heat-labile enterotoxin type I. Microbiol. Immunol. *40*, 71–75.

Simmons, C.P., Ghaem-Magami, M., Petrovska, L., Lopes, L., Chain, B.M., Williams, N.A., and Dougan, G. (2001). Immunomodulation using bacterial enterotoxins. Scand. J. Immunol. *53*, 218–226.

Sixma, T.K., Kalk, K.H., van Zanten, B.A., Dauter, Z., Kingma, J., Witholt, B., and Hol, W.G. (1993). Refined structure of Escherichia coli heat-labile enterotoxin, a close relative of cholera toxin. J. Mol. Biol. *230*, 890–918.

Sreenivasan, U., and Axelsson, P.H. (1992). Buried water in homologous serine proteases. Biochemistry *31*, 12785–12791.

Sun, J.B., Holmgren, J., and Czerkinsky, C. (1994). Cholera toxin B subunit: an efficient transmucosal carrier-delivery system for induction of peripheral immunological tolerance. Proc. Natl. Acad. Sci. USA *91*, 10795–10799.

Swaminathan, C.P., Surolia, N., and Surolia, A. (1998). Role of water in the specific binding of mannose and mannooligosaccharides to concanavalin A. J. Am. Chem. Soc. *120*, 5153–5159.

Svennerholm, A.M., and Holmgren, J. (1978). Identification of *Escherichia coli* heat-labile enterotoxin by means of a ganglioside immunosorbent assay (GM1-ELISA) procedure. Curr. Microbiol. *1*, 19–23.

Teneberg, S., Hirst, T.R., Ångström, J., and Karlsson, K.A. (1994). Comparison of the glycolipid-binding specificities of cholera toxin and porcine *Escherichia coli* heat-labile enterotoxin: identification of a receptor-active non-ganglioside glycolipid for the heat-labile toxin in infant rabbit small intestine. Glycoconj. J. *11*, 533–540.

Teneberg, S., Berntsson, A., and Ångström, J. (2000). Common architecture of the primary galactose binding sites of *Erythrina corallodendron* lectin and heat-labile enterotoxin from *Escherichia coli* in

relation to the binding of branched neolactohexaosylceramide. J. Biochem. (Tokyo) *128*, 481–491.

Teneberg, S., Alsén, B., Ångström, J., Winter, H.C., and Goldstein, I.J. (2003). Studies on Galα3-binding proteins: comparison of the glycosphingolipid binding specificities of *Marasmius oreades* lectin and *Euonymus europaea* lectin. Glycobiology *13*, 479–486.

Toone, E.J. (1994). Structure and energetics of protein-carbohydrate complexes. Curr. Opin. Struc. Biol. *4*, 719–728.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. *30*, 1022–1025.

Williams, S.J., and Davies, G.J. (2001). Protein-carbohydrate interactions: learning lessons from nature. Trends Biotechnol. *19*, 356–362.

Yang, H.J., and Hakomori, S.I. (1971). A sphingolipid having a novel type of ceramide and lacto-N-fucopentaose 3. J. Biol. Chem. *246*, 1192–1200.

Zhang, Z., Merritt, E.A., Ahn, M., Roach, C., Hou, Z., Verlinde, C.L., Hol, W.G., and Fan, E. (2002). Solution and crystallographic studies of branched multivalent ligands that inhibit the receptor-binding of cholera toxin. J. Am. Chem. Soc. *124*, 12991–12998.

Accession Numbers

Atomic coordinates and structure factors have been deposited with the PDB under accession code 1TL0.